DOCUMENT RESUME

ED 369 816 TM 021 455

AUTHOR Han, Lei; Hoover, H. D.
TITLE Gender Differences in Achievement Test Scores.
PUB DATE Apr 94
NOTE 25p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New Orleans, LA, April 5-7, 1994).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Academic Achievement; *Achievement Tests; Age Differences; Educational Trends; Elementary School Students; Elementary Secondary Education; Language Proficiency; Mathematics Achievement; Research Needs; *Scores; Secondary School Students; *Sex Differences; *Standardized Tests; Test Results; Trend Analysis
IDENTIFIERS Iowa Tests of Basic Skills; Iowa Tests of Educational Development; *Standardization; Tests of Achievement and Proficiency

ABSTRACT

National standardization data were used to reexamine the evidence of gender differences in achievement test scores reported in previous studies. Changes in differences over time, from 1963 to 1992, were examined, as were differences across all grade and achievement levels. Subjects participated in Iowa Tests of Basic Skills (ITBS), Iowa Tests of Educational Development (ITED), and Tests of Achievement and Proficiency (TAP) national standardizations from 1963 to 1992. Results generally agree with previous studies. In general, the nature and magnitude of differences in achievement were similar over the years. Average differences in achievement between males and females were relatively small in most subject areas, with females scoring higher in reading, language skills, and mathematics computation up to age 15 when the opposite is true in mathematics. Females generally scored higher in verbal tests at the upper grade levels. Male performance was generally more variable than that of females, and these differences in variability lead to interaction between ability level and differences in achievement for males. Males at average and below-average abilities have markedly lower achievement than females. Suggestion are made for future research into language deficits of low-achieving boys. Five tables present study results. (Contains 15 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 369 816

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

LEI HAN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Gender Differences in Achievement Test Scores

Lei Han
H. D. Hoover
The University of Iowa

Paper presented at the Annual Meeting of the
National Council on Measurement in Education
New Orleans, LA
April, 1994

TM021455

Gender Differences in Achievement Test Scores

Much research on gender differences in academic performance has been conducted since the first extensive review of the literature by Maccoby and Jacklin (1974). The issues of gender differences have continued to attract the attention of educators and the public in recent years. Wilder and Powell (1989) summarized over 200 research studies, most of them published after 1980, and drew conclusions similar to those of Maccoby and Jacklin. In tests of mathematical or quantitative ability, males tended to do better than females, particularly in secondary school and beyond; females tended to do slightly better than males in many tests of verbal skills (particularly writing). In addition, Wilder and Powell found evidence that gender gaps have been narrowing over the years. However, findings regarding the magnitude of gender differences were inconsistent across studies.

Some research based on meta-analyses has concluded that the gender differences in verbal and mathematics abilities are negligible or virtually nonexistent in the general population (Hyde and Linn, 1988; Hyde, Fennema, and Lamon, 1990; Hyde, 1991). Some other meta-analyses (Hengkietisak & Cleary, 1989; Cleary, 1991) based on 18 widely used standardized achievement and aptitude test norm data sets have concluded that gender differences are larger in selected samples in which students choose to take the tests (e.g. SAT, ACT and GRE test takers) than in unselected samples in which the entire class of students take the tests (e.g. ITBS test takers). In general, females score lower than males on most tests in the selected samples, and these differences increase with age, task complexity, and the quantitative nature of the tasks. Differences in unselected samples are not as indicative of higher male performance. There also are interactions among these factors and gender gaps. For mathematics tests, females have slightly higher test scores than males for the young age group, but the opposite is true for the age group beyond fifteen. The average effect size is greater than 0.3 for PSAT, SAT, ACT test takers in favor of males. Females outperform males in verbal test scores with an effect size (d) of

about 0.20 for unselected samples. Females have the greatest advantage on spelling tasks (d= 0.40); however, this advantage decreases as the complexity of tasks increase (e.g. reading comprehension d= 0.2)

Two extensive longitudinal studies (Martin & Hoover, 1987; Becker & Forsyth, 1990) conducted on a statewide sample of third through twelfth grade students in Iowa who completed the Iowa Tests of Basic Skills (ITBS), or the Iowa Tests of Educational Development (ITED) indicated that on almost all subtests of the ITBS and ITED, male scores were more variable than female scores. On most subtests there was an interaction between difference in performance and achievement level. For instance, in those tests where females had higher average scores than males, the largest differences favoring females tended to be found below the median of the distribution, while in those tests where males had higher average scores, the largest differences tended to occur above the median. The authors suggested that to gain a better understanding of gender differences in achievement, it would be more informative to compare entire distributions for males and females rather than group averages.

The present study was designed to use national standardization data to reexamine the evidence of gender differences reported in the previous studies. Our interests were not only in general verbal and mathematics abilities but also specific components of these abilities such as math computation, math concepts and estimation, and math problem solving. In addition, other areas of the school curriculum such as language, the use of sources of information, social studies and science were studied. Three types of differences were examined: changes in differences over time from 1963 to 1992; differences across all grade levels from kindergarten to high school; and differences across achievement levels. In addition, the score variability for males and females was compared.

Method

Subjects

The subjects in this study were those students who participated in various ITBS, ITED, and the Tests of Achievement and Proficiency (TAP) national standardizations from 1963 to 1992. The ITBS has been nationally standardized approximately every seven years from 1963 to 1992. In each of these five standardizations about 15,000 students in each grade from grade three through grade eight (grades K-2 were added in 1977) were carefully selected to form a nationally representative sample. The TAP was standardized along with the ITBS in the 1963 to 1984 standardizations in grade nine through twelve but no gender differences analyses were performed. In the 1992 standardization the ITED was added in grades nine to twelve.

Instruments

The achievement tests used in this study are the Iowa Tests of Basic Skills (Lindquist and Hieronymus, 1964; Hieronymus et al., 1970, 1978, 1986; Hoover et al., 1993) for kindergarten through grade 8, the Iowa Tests of Educational Development (Feldt, et al., 1993) for grade 9 through grade 12, and the Tests of Achievement and Proficiency (Scannell, et al.,1993). The ITBS, a multi-level achievement battery consists of five to thirteen subtests depending on test level, three subscale composites, a battery composite based on reading, language, and mathematics tests (core total), and the total battery composite. The ITED consists of seven subtests, core total, and the total battery composite. The TAP consists of eight subtests, core total, and total battery composite. A summary of these subtests and composites is listed in table 1. Tests in the same row of this table are on a common scale across the three batteries. Mathematics computation is an optional test for the ITBS and TAP and is therefore listed at the end of the table, along with a total mathematics score including computation.

Analyses

Since gender difference in grade equivalent (GE) scores for previous standardizations from 1963 to 1984 have been reported in ITBS manuals (Hieronymus and Hoover, 1986), data for the 1992 standardization were analyzed in the same manner to examine the changes of gender differences over the last 30 years. Differences in performance at selected percentiles (10th, 50th, and 90th) for males and females were examined to study the possible interaction between achievement level and gender. For the 1992 standardization data standard score descriptive statistics for each subtest and composite were computed for males and females for grade 4, 8, 11 using the ITBS, ITED, and TAP standardization samples. Effect sizes were computed by subtracting males score from females score at selected percentile ranks and dividing by the average standard deviation of males and females. The ratio of the male standard deviation to the female standard deviation was computed to describe the relative variability of the two groups. In addition, comparisons of performance at the 10th, 50th, and 90th percentiles were made. Inferential statistical tests were not reported because the standard errors associated with difference in performance between males and females in the ITBS, ITED, and TAP standardization were so small that almost all observed differences would be statistically significant.

Results and Discussion

Comparison of Gender Differences over the Years

Table 2 reports longitudinal comparison of gender differences in terms of grade equivalent (GE) scores for the 1963, 1970, 1977, 1984 and 1992 ITBS national standardizations for kindergarten though grade 8. These differences are summarized by total test area scores at the 10th, 50th, and 90th percentiles. Positive values indicate females have higher scores. In general, the nature and magnitude of the differences between male and female scores have remained similar over the last 30 years. Females have consistently scored higher than males in Reading Comprehension and Language Total. These differences are especially pronounced at average (PR=50) and below average

(PR=10) achievement levels. In vocabulary, work study (referred to as sources of information in the 1992 edition), and mathematics, the differences have not in general been as large, but females in most cases still have higher scores than males at the middle and lower achievement levels. At the upper achievement level (PR=90) males scores have tended to equal or slightly surpass those of females in these test areas. There does appear to have been a slight increase in the relative superiority of females over this period of time in Reading Comprehension at all achievement levels. An apparent similar change in relative performance on the Mathematics Total score is at least partially explained by the addition in 1977 of a computation test to the test battery. As will be apparent when individual test scores are discussed later, computation is the area of mathematics in which females consistently score higher than males at all achievement levels. Including computation in the total math score, along with the math concepts and math problem solving tests, makes female to male comparisons more favorable toward females. Table 3 reports gender differences for the 1984 and 1992 standardizations with and without computation included in the total math score. As can be seen, including computation in the total math score does increase the performance of females relative to that of males.

Comparison of Group Means

In order to simplify a more detailed study of gender differences on the 1992 national standardization, we focused on students at selected grade levels instead of every grade. Table 4 reports descriptive statistics for grade 4 and grade 8 students tested with ITBS, grade 11 students tested with ITED, and grade 11 students tested with TAP on each subtest and composite scores by gender. Mean effect sizes and the ratio of male and female standard deviations are also reported in this table. As can be seen, in general these effect sizes were small. With the exception of the language tests, few were greater the one-quarter of a standard deviation. The relatively large effect sizes on the language tests were all in favor of females. Females scored consistently better than males on Spelling, Capitalization, Punctuation, and Language Usage. Because the results were roughly

comparable across the language subtests in grade 4 and 8, the results from Language Total provide an adequate summary of the gender differences in language skills. Language Total on the ITBS is also comparable in content to the individual language test on the ITED and TAP at grade 11. There was a tendency for the differences in language to increase across grade with the largest effect size of 0.53 at grade 11 on the TAP.

The results for Mathematics Total showed that there was little difference in the average performance of males and females. Focusing on total math, as was mentioned earlier, tends to obscure the pattern of results from specific subtests and at different score levels. At grade 4 and grade 8, males were slightly better than females on Mathematics Concepts. Differences in Mathematics Problem Solving at the mean level were minimal. Females on the other hand, were better than males in Mathematics Computation, and the differences were larger for grade 8 than for grade 4. Math computation is not included in the total math score for these and subsequent comparisons for a couple of reasons. First, the ITED does not contain any computation items. Second, the National Council of Teachers of Mathematics is strongly encouraging schools to de-emphasize rote computational procedures. On the ITBS and TAP computation is still included in the batteries but as an optional subtest. For those interested in results with computation, the last two columns of Tables 4 and 5 present results for math computation and a total math score including computation. In general, the mean differences for males and females on the mathematics subtests and total scores with or without computation were quite small across grades. However, these results are misleading to some extent. The interaction between gender differences and achievement level will be discussed later.

Mean differences on Social Studies and Science were relatively small at all grade levels. The largest differences were ones favoring females on the ITED at grade 11. A possible explanation for this result will be given later. The results for Sources of Information Total were similar to those on Mathematics Total at grades 4 and 8. Females had higher means on Reference Materials, while males had higher means on Maps and

Diagrams. Because the differences were greater on the Reference Materials test, females had somewhat higher total scores. Differences in male and female performance on the ITED and TAP sources of information scores are larger than for the ITBS. This test is very similar to the ITBS Reference Materials on the ITED and even though TAP does contain some maps it is still weighted in content toward reference materials. The larger effect sizes in favor of females at grade 11 reflect these content differences between ITBS and ITED/TAP on sources of information.

In most cases the mean differences between males and females on ITED and TAP subtests were larger than on ITBS, and were in favor of females except for Mathematics Total on the ITED. The largest differences occurred on Language Total, Literary Materials, and Sources of Information. In most cases the differences favoring females were also somewhat larger on the ITED than on the TAP. The ITED relies heavily on text based stimulus materials in both the Analysis of Social Studies Materials and Analysis of Science Materials tests. It also contains a test titled the Ability to Interpret Literary Materials. These tests all require a great deal of reading. Females consistently score higher than males across all grade levels in Reading Comprehension.

Comparison of score variability

One aspect of gender differences that was most noticeable in the Martin and Hoover (1987) and Becker and Forsyth (1990) studies was greater variability in almost every test area for males than for females. Both of these studies examined performance in a large longitudinal sample in Iowa. In this study, test score variability for a representative sample of students nationally was examined. The ratio of the male to female standard deviations is listed in Table 4 for each subtest. These ratios were all greater than or equal to 1.00 except for Reading Comprehension in grade 3. The standard deviations for males and females were very similar for the reading, language, and references subtests in grade 4 and grade 8. For other test areas, males exhibited greater variability than females. The variability ratios

tended to increase across grades. In grade 11 the male standard deviation is about 15 percent greater than the female standard deviation on most tests for both ITED and TAP.

Comparison of Group Difference by Ability Level

The differences in gender group variation make it apparent that comparing means does not provide a clear picture of the relationship between gender and achievement. In order to examine the interaction between gender differences and achievement level, effect sizes on each subtest at the 10th, 50th and 90th percentiles were computed and are listed in Table 5. The effect sizes on Vocabulary, Reading Comprehension, and Reading Total were relatively small at all achievement levels for both grades 4 and 8. Moderate effect sizes favoring females were obtained on Reading Comprehension for TAP and Content Area Reading for the ITED at grade 11 (this score is based on reading comprehension types of items contained in the Analysis of Natural Science Materials, Analysis of Social Studies Materials, and Ability to Interpret Literary Materials subtests of the ITED) . The effect sizes for reading were substantial at low achievement levels (e.g. 0.62 on the ITED).

The language tests are the only ones in which the differences in achievement are relatively similar across ability levels. Females outperformed males at all levels. The effect sizes were somewhat larger at the upper grade levels and at the 10th and 50th percentiles.

The results from the mathematics subtests demonstrate why comparisons of means are sometimes misleading. Males outperformed females on Mathematics Concepts and Mathematics Problems Solving, particularly at upper percentile rank levels, while females were slightly better than males at low percentile rank levels. As a result the differences at the mean were quite small. In Mathematics Computation, the differences were small at the upper end of distribution. The advantage for females was most pronounced at low achievement levels. The results for Mathematics Total at all grade levels, with or without computation, again showed a clear effect size by achievement level interaction. A slight advantage for above average males and a similar advantage for below average females is shown.

The patterns of differences across percentile levels for Social Studies, Science, and Maps and Diagrams were very similar to that of Mathematics Total for grade 4 and grade 8. Males outperformed females at upper achievement levels, but the effect sizes were quite small at the median and below. At the high school level, males outperformed females at upper achievement levels in both Social Studies and Science. Similar differences in favor of females were observed at low achievement levels.

The results for References Materials were similar to that of the language tests where females outperformed males at all achievement levels. The effect sizes at the median on ITBS Sources of Information Total show no noticeable differences between males and females. This result is somewhat misleading for two reasons: First, this total score is an average of Maps and Diagrams and References Materials which exhibit opposite gender differences. Second, the differences are quite different for above and below average students. On the sources tests for TAP and ITED, there were no differences for high scoring students, but, females showed substantial advantages at and below the median.

In summary, a comparison of gender differences across achievement levels showed that females had a definite performance advantage on the language tests and Reference Materials at all achievement levels. However, in other test areas in which females showed superior performance, the greatest differences occurred below the medians of the distribution. When males had better scores, the largest differences occurred above the median. These results demonstrate that differences at the mean may provide misleading information. A better understanding of gender differences in achievement was obtained from comparing the entire score distributions for males and females.

Summary

The results from this study generally agree with the results of previous studies on gender differences in unselected populations of K-12 students; e.g., Martin and Hoover (1987) and Becker and Forsyth (1990). Average differences in achievement between males and females were relatively small in most subject areas. Females generally scored higher than

males in reading, language skills, and mathematics computation, and the magnitude of these differences tended to be somewhat larger at upper grade levels. The results also support the position of Martin and Hoover that differences in mean performance may provide somewhat misleading information regarding gender differences in achievement. In general male performance was more variable than that of females and this effect was most pronounced at upper grade levels. These differences in variability lead to an interaction between ability level and differences in achievement between males and females. Much of the recent literature on gender differences focuses on data from samples in which students choose, or select, to take the tests (e.g., National Merit, SAT, ACT, and GRE). Since most of these selected samples come from the upper end of the ability distribution, it is probably not surprising that a perception that males score higher than females on achievement and aptitude tests is prevalent. In addition, the focus of much of the research on these self-selected samples has been on performance in math and science. These two areas had the largest effect sizes favoring males in the present study for above average (90th percentile) students.

By far the largest effect sizes in this study, and in the previous ITBS standardization studies summarized in Table 2, are the result of markedly lower achievement of males at average and below average achievement levels. These differences are especially large in reading and language. They also increase in magnitude in all test areas as students get older.

The following quote from Hoover and Martin's longitudinal study of Grade 3 through Grade 8 students would still seem to hold.

" In view of the importance of reading and language skills in further academic achievement, it would not seem unwarranted to suggest that some of the emphasis in future research on the nature and causes of sex differences in achievement be focused on the language deficits of low achieving boys."

In light of the results from this study, extending this suggestion to include deficits of low achieving teen-age males in all subject matter areas would not seen unreasonable.

Reference

Becker, D. F., & Forsyth, R. A. (1990). Gender differences in academic achievement in grades 3 through 12: a longitudinal analysis. Paper presented at the annual meeting of the America Educational Research Association, Boston, MA.

Cleary, T. A. (1991). Gender differences in aptitude and achievement scores. Paper present at the fifty-second ETS Invitational Conference. New York city, New York.

Feldt, L. S., Forsyth, R. A., Ansley, T.N., and Alnot, S.D. (1993). Iowa tests of educational development, forms k and l. Chicago: Riverside.

Hengkietisak, S. & Cleary, T. A. (1989). Gender differences in mathematics ability: A meta-analysis of mathematics aptitude and achievement test scores. Paper presented at annual meeting of America Psychological Association. New Orleans, LA.

Hieronymus, A. N., Hoover, H. D. (1986). Iowa tests of basic skills, forms G and H. Chicago: Riverside.

Hieronymus, A. N., Lindquist, E. F. (1970). Iowa tests of basic skills, forms 5 and 6. Boston: Houghton Mifflin Company.

Hieronymus, A. N., Lindquist, E. F., and Hoover, H. D. (1978). Iowa tests of basic skills, forms 7 and 8. Chicago: Riverside.

Hoover, H. D., Hieronymus, A. N., Frisbie, D. A., and Dunbar, S. B. (1993). Iowa tests of basic skills, forms k and l. Chicago: Riverside.

Hyde, J. S. & Linn, M. C. (1988). A meta-analysis of gender differences in verbal abilities. Psychological Bulletin, 104, 53-69.

Hyde, J. S., Fennema, E. & Lamon, S. (1990). Gender differences in mathematics performance: a meta-analysis. Psychological Bulletin, 107, 139-155.

Hyde, J. S. (1991). Gender and sex: so what has meta-analysis done for me? Paper presented at American Psychological Association Convention, San Francisco, CA.

Lindquist, E. F., Hieronymus, A. N. (1964). Iowa tests of basic skills, forms 3 and 4. Boston: Houghton Mifflin Company.

Martin, D. J., & Hoover, H. D. (1987). Sex differences in educational achievement: a longitudinal study. Journal of Early Adolescence, 7, 65-83.

Scannel, D. P., Haugh, O., Loyd, B.H., and Risinger, C.F. (1993). Tests of achievement and proficiency, forms k and l. Chicago: Riverside.

Wilder, G. Z., & Powell, K. (1989). Sex differences in test performance: a survey of the literature. (College Board Report. No. 89-3). New York: College Entrance Examination Board.

Table 1.
ITBS, ITED, and TAP Subtests and Composites

	ITBS	TAP	ITED
RV	Vocabulary	Vocabulary	Vocabulary
RC	Reading Comprehension	Reading Comprehension	Content Area Reading
RT	Reading Total	Reading Total	Reading Total
L1	Listening		
L2	Capitalization		
L3	Punctuation		
L4	Language/Usage & Expression		
LT	Language Total	Written Expression	Correctness and Appropriateness of Expression
M1	Mathematics Concepts & Estimation		
M2	Math Problem Solving & Data Interpretation		
MT	Mathematics Total	Math Concepts & Problem Solving	Ability to Do Quantitative Thinking
CT	**Core Total**	**Core Total**	**Core Total**
LM			Ability to Interpret Literary Materials
SS	Social Studies	Social Studies	Analysis of Social Studies Materials
SC	Science	Science	Analysis of Science Materials
S1	Maps and Diagrams		
S2	Reference Materials		
ST	Sources of Information Total	Information Processing	Use of Sources of Information
CC	**Complete Composite**	**Complete Composite**	**Complete Composite**
M3	Mathematics Computation	Mathematics Computation	
MTX	Mathematics Total (including Computation)	Mathematics Total (including Computation)	

Table 2
Iowa Tests of Basic Skills Longitudinal Comparison of Gender Differences in Achievement 1963, 1970, 1977, 1984, and1992.

		Vocabulary					Reading					Total Language					Total Work Study					Total Mathematics					Composite				
PR	Grade	'63	'70	'77	'84	'92	'63	'70	'77	'84	'92	'63	'70	'77	'84	'92	'63	'70	'77	'84	'92	'63	'70	'77	'84	'92	'63	'70	'77	'84	'92
	8	0	0	-4	-1	-2	-1	0	2	1	1	6	7	4	5	5	-1	-1	0	-1	-2	-2	-2	0	-1	-3	0	1	1	0	0
	7	0	0	-3	-3	-1	1	0	1	1	3	9	6	5	5	6	1	0	0	-1	0	2	-2	0	-1	-1	3	1	1	0	1
	6	0	0	-2	-2	-2	1	1	1	1	3	6	6	5	5	4	1	0	0	0	-1	1	-1	0	-1	-1	1	1	1	1	1
	5	0	1	-1	-1	-3	1	2	0	0	3	4	5	5	4	3	-1	0	-1	-1	-1	1	-1	0	-1	-1	1	2	1	0	0
90	4	0	1	-1	-1	-3	0	3	0	0	2	5	5	5	3	2	-1	0	-1	-1	-2	1	-1	0	-1	-2	1	2	1	0	0
	3	0	2	0	1	0	2	3	1	2	2	5	4	4	3	2	1	0	0	0	-1	1	1	0	0	-1	1	2	0	1	0
	2			0	0	0			1	2	2			4	3	2			0	1	-2			0	-1	-2			1	2	0
	1			0	-1	0			2	2	2			2	0	3					0			-1	0	0			1	1	1
	k			-1	-1	0					0			0	1	2								-1	0	0			0	0	1
	8	0	1	-1	-2	-1	1	4	6	6	4	9	11	11	9	9	2	4	6	5	4	0	1	5	1	0	2	4	5	4	3
	7	0	1	0	-1	0	2	4	5	4	3	10	11	10	8	10	1	3	4	3	4	1	1	4	2	2	3	5	5	3	3
	6	2	2	1	-1	0	3	4	4	3	4	8	10	9	6	7	3	2	3	3	3	2	1	3	2	2	4	4	4	2	2
	5	1	2	1	0	-1	2	3	6	3	4	7	7	8	5	5	1	2	3	2	1	1	1	2	1	1	3	3	3	2	2
50	4	1	2	1	1	-1	2	3	2	2	2	6	6	7	4	4	1	2	2	2	1	2	1	1	1	0	3	3	3	2	1
	3	2	2	1	1	0	2	3	2	2	2	5	4	5	4	4	2	2	2	1	1	1	1	1	0	0	3	3	3	2	1
	2			1	1	0			2	2	1			5	3	3			2	1	0			1	0	0			2	1	1
	1			1	0	0			2	1	3			3	0	2					1			0	0	0			1	0	1
	k			0	0	1					1			1	1	1								0	1	0			0	1	1
	8	2	3	3	-1	1	2	4	6	8	3	6	6	8	9	10	1	3	5	6	4	0	1	4	3	3	2	2	5	5	3
	7	0	5	3	1	2	1	5	5	6	6	5	6	7	7	8	0	3	4	3	4	0	1	3	3	4	1	2	4	4	4
	6	2	3	2	0	2	3	2	3	4	6	6	4	6	6	7	3	3	4	3	3	1	1	2	2	5	2	2	3	3	4
	5	2	3	2	1	-1	2	2	2	4	5	5	3	5	4	7	1	1	3	3	4	1	1	2	1	3	1	1	3	2	3
10	4	1	2	2	1	0	2	1	2	3	2	2	3	4	3	4	1	1	3	3	3	1	1	2	1	2	1	1	2	3	2
	3	1	1	1	1	1	1	1	1	3	2	2	1	3	2	3	1	1	2	2	2	1	1	1	1	1	1	1	2	2	2
	2			1	1	1			1	2	1			2	1	2			2	1	1			1	0	1			1	1	1
	1			0	0	0			1	1	1			1	0	1					1			0	0	1			1	0	1
	k			0	0	1					0			1	0	1								0	0	0			0	0	1

*Entry values are in GE units-months
*Positive differences favor females

Table 3
Iowa Tests of Basic Skills
Gender Differences on Mathematics Total with and without Computation

Grade	1984		1992	
	Without	With	Without	With
8	-3	-1	-5	-3
7	-2	-1	-4	-1
6	-2	-1	-1	-1
5	-2	-1	-1	-1
4	-1	-1	-3	-2
3	-1	0	-2	-1
8	0	1	-1	0
7	1	2	1	2
6	1	2	1	2
5	0	1	0	1
4	1	1	-1	0
3	0	0	-1	0
8	1	3	1	3
7	2	3	4	4
6	1	2	4	5
5	0	1	2	3
4	1	1	2	2
3	1	1	1	1

* Entry values are in GE units-months
* Positive values favor females

Table 4
ITBS, ITED, and TAP Descriptive Statistics, Spring 1992 National Standardization

Battery			RV	RC	RT	L1	L2	L3	L4	LT	M1	M2	MT	CT	LM	SS	SC	S1	S2	ST	CC	M3	MTX
ITBS	Male	Mean*	200.8	200.7	200.8	199.3	198.5	200.5	201.3	200.0	201.5	204.1	202.8	**201.5**		204.5	205.6	205.9	198.4	202.2	**203.2**	199.8	201.9
Grade 4		SD	23.5	28.3	24.4	24.6	35.3	33.6	33.7	27.5	23.3	29.5	25.0	**23.4**		27.6	30.7	32.5	22.4	25.7	**23.9**	21.1	22.2
	Female	Mean	199.3	205.5	202.5	206.3	209.1	208.5	208.0	208.1	199.6	202.6	201.1	**204.1**		202.0	202.7	203.0	202.9	203.0	**203.6**	202.7	201.8
		SD	22.4	28.6	24.0	24.6	35.1	33.6	33.9	27.2	20.9	26.8	22.4	**22.5**		24.6	27.1	28.5	21.7	23.2	**22.0**	19.8	19.8
		d**	**-0.07**	**0.17**	**0.07**	**0.28**	**0.30**	**0.24**	**0.20**	**0.30**	**-0.09**	**-0.05**	**-0.07**	**0.11**		**-0.10**	**-0.10**	**-0.10**	**0.21**	**0.03**	**0.02**	**0.14**	**0.00**
		R***	**1.05**	**0.99**	**1.02**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.12**	**1.10**	**1.17**	**1.04**		**1.12**	**1.13**	**1.14**	**1.03**	**1.11**	**1.09**	**1.07**	**1.12**
ITBS	Male	Mean	245.9	245.6	247.9	242.8	243.2	240.8	242.2	242.6	251.8	251.1	251.7	**248.4**		252.4	252.3	254.6	246.1	250.6	**251.4**	246.8	250.4
Grade 8		SD	31.4	41.9	34.5	35.5	50.1	50.8	52.2	41.4	35.0	43.7	37.3	**34.5**		44.1	45.8	46.6	40.5	40.9	**36.3**	37.6	34.9
	Female	Mean	248.3	252.0	250.4	259.4	259.3	261.6	259.9	260.4	248.5	249.3	249.1	**253.8**		249.9	248.9	249.6	258.8	254.3	**253.3**	256.8	252.0
		SD	29.8	40.6	33.3	34.4	49.5	49.8	50.6	40.1	31.6	40.1	33.7	**32.8**		39.7	39.3	42.7	38.5	37.8	**33.2**	35.7	31.5
		d	**-0.04**	**0.15**	**0.07**	**0.47**	**0.32**	**0.41**	**0.34**	**0.44**	**-0.10**	**-0.04**	**-0.07**	**0.16**		**-0.06**	**-0.08**	**-0.11**	**0.32**	**0.09**	**0.05**	**0.27**	**0.05**
		R	**1.06**	**1.03**	**1.04**	**1.03**	**1.01**	**1.02**	**1.03**	**1.03**	**1.11**	**1.09**	**1.11**	**1.05**		**1.11**	**1.16**	**1.09**	**1.05**	**1.08**	**1.09**	**1.05**	**1.11**
ITED	Male	Mean	271.9	269.9	271.1					268.1			279.2	**273.7**	266.2	271.9	272.9			269.7	**272.5**		
Grade 11		SD	38.3	49.3	41.0					48.1			43.4	**40.1**	44.1	49.5	49.7			47.1	**39.4**		
	Female	Mean	276.7	284.1	280.6					289.9			275.0	**282.3**	284.9	281.1	280.9			282.9	**282.2**		
		SD	32.9	42.3	35.2					43.8			38.2	**35.3**	38.0	44.0	43.2			40.6	**34.6**		
		d	**0.13**	**0.31**	**0.25**					**0.47**			**-0.10**	**0.23**	**0.45**	**0.20**	**0.17**			**0.30**	**0.26**		
		R	**1.16**	**1.16**	**1.16**					**1.10**			**1.14**	**1.14**	**1.16**	**1.13**	**1.15**			**1.16**	**1.14**		
TAP 11	Male	Mean	272.4	267.3	270.0					261.2			273.7	**268.8**		274.5	273.1			269.5	**271.8**	270.1	272.4
		SD	38.1	49.9	41.5					45.7			42.8	**39.6**		52.7	52.9			49.5	**42.6**	44.7	41.5
	Female	Mean	273.6	276.3	275.0					284.7			274.2	**278.3**		274.8	274.2			280.6	**278.3**	277.4	275.8
		SD	33.1	44.8	36.7					42.8			37.4	**35.4**		44.8	43.6			42.5	**36.7**	39.2	36.2
		d	**0.03**	**0.19**	**0.13**					**0.53**			**0.01**	**0.25**		**0.01**	**0.02**			**0.24**	**0.16**	**0.17**	**0.09**
		R	**1.15**	**1.11**	**1.13**					**1.07**			**1.14**	**1.12**		**1.18**	**1.22**			**1.16**	**1.16**	**1.14**	**1.14**

* Means and SD's are in standard score units
** d=(female mean - male mean)/ average SD of male and female
*** R=male SD/female SD

Table 5
Gender Differences by Ability Level on ITBS, ITED, and TAP

Grade	PR	RV	RC	RT	L1	L2	L3	L4	LT	M1	M2	MT	CT	LM	SS	SC	S1	S2	ST	CC	M3	MTX
	90	-0.12	0.17	0.04	0.31	0.38	0.20	0.18	0.29	-0.26	-0.19	-0.19	**0.07**		-0.26	-0.29	-0.28	0.19	-0.12	**-0.10**	0.03	-0.16
ITBS 4	50	0.05	0.14	0.06	0.25	0.25	0.23	0.20	0.30	-0.11	-0.06	-0.09	**0.12**		-0.14	-0.07	-0.09	0.21	0.00	**0.01**	0.13	-0.02
	10	0.00	0.12	0.07	0.31	0.32	0.20	0.20	0.28	0.09	0.10	0.10	**0.17**		0.10	0.07	0.15	0.24	0.21	**0.15**	0.30	0.16
	90	-0.05	0.11	0.03	0.41	0.23	0.27	0.26	0.33	-0.21	-0.18	-0.24	**0.06**		-0.21	-0.35	-0.27	0.17	-0.07	**-0.11**	0.12	-0.16
ITBS 8	50	-0.07	0.19	0.10	0.55	0.38	0.51	0.43	0.52	-0.10	-0.01	-0.05	**0.24**		-0.06	-0.01	-0.09	0.40	0.03	**0.12**	0.34	0.08
	10	-0.07	0.18	0.08	0.50	0.28	0.36	0.37	0.40	0.04	0.08	0.05	**0.19**		0.09	0.09	0.02	0.32	0.14	**0.15**	0.29	0.15
	90	-0.06	0.02	0.00					0.25			-0.31	**-0.01**	0.00	-0.06	-0.10			-0.03	**-0.03**		
ITED 11	50	0.07	0.34	0.25					0.59			-0.08	**0.35**	0.47	0.28	0.21			0.43	**0.36**		
	10	0.53	0.62	0.46					0.53			0.06	**0.35**	0.87	0.28	0.30			0.40	**0.36**		
	90	-0.09	0.05	-0.02					0.41			-0.19	**0.07**		-0.23	-0.24			-0.04	**-0.08**	-0.03	-0.11
TAP 11	50	-0.06	0.20	0.09					0.64			0.00	**0.29**		0.00	0.03			0.29	**0.20**	0.17	0.12
	10	0.29	0.39	0.32					0.55			0.18	**0.36**		0.26	0.32			0.38	**0.35**	0.43	0.24

* Eentry value = (female score - male score)/ average SD of male and female.